2018 4th International Conference on Education & Training, Management and Humanities Science (ETMHS 2018)

Research on Application of Linear Frequency Modulation Chirp Z in Pitch Piano Detection

Yao Fang

Department of music of Taiyuan Normal University, Taiyuan, Shanxi, China, 030619

Keywords: linear frequency modulation; Chirp Z; simulation; pitch detection

Abstract: Through theoretical analysis, it is confirmed that the linear frequency modulation Z transform is effective in refining the single frequency signal. A more accurate frequency value can be obtained by finding the position of the extreme point. The pitch detection requirements of the piano and the nature of the linear frequency modulation Z transform are analyzed and used in the pitch detection to make the detection accuracy, reliability, anti-interference, real-time and stability meet the requirements of the factory. The labor intensity of workers is greatly reduced, the production efficiency of the factory is improved, and the quality of the products is more stable and reliable.

1. Introduction

Pitch detection is a hot and difficult problem in speech research. Pitch detection in musical instruments has always been a technical problem puzzling musical instrument making factories. Even the pitch detection of piano produced by some piano factories is entirely realized by human ears. So the accuracy and reliability of the test results are closely related to the state and experience of people. It is not only a job with high labor intensity for the workers, but also a large amount of money to train new workers every year for the factory, because each worker is limited in the time of tuning, usually for 3 years, and if the time is too long, the hearing will be insensitive. Therefore, how to detect the pitch of the piano quickly and accurately is a combination of theoretical research and practical application. At present, autocorrelation function method and short-time mean amplitude difference function method are used to detect the pitch in the theoretical field, and in order to improve the reliability of detection, central clipping and low-pass filtering preprocessing and post-processing smoothing techniques are also used. However, the above methods are not suitable for Piano pitch detection. Because the accuracy of the above methods is not high, the object of the study is speech. The piano's pitch detection is very demanding, and the object of study is the musical sound, which must be accurate to one minute. The score is in logarithmic relationship with the pitch of the piano. The first string has a pitch of 0.02 Hz, and the error of every string must be within 3~4 tones. It can be seen that special methods must be adopted.

2. Linear Frequency Modulation Function Transformation

Fourier transform is a mathematical tool linking time domain and frequency domain. Many improvements have been made to Fourier transform, and various improved algorithms have been proposed. Among them, the CZT transformation is better in accuracy. The number of input points and output points of the transform may not be equal, so as to achieve the goal of "thinning" in frequency domain. The name of the transform comes from radar professional vocabulary, that is, "linear frequency modulation signal", also known as chirp signal.

2.1 The principle of chirp transformation.

Linear frequency modulation Z transform can highlight the frequency range we want to study. Let $x[n]$ be a N point sequence and $X(e^{ix})$ represents its Fourier transform. The M samples of $X(e^{iw})$ are calculated, and these samples are arranged at equal angle intervals on the unit circle, as shown in

Figure 1, that is, at the following frequencies:

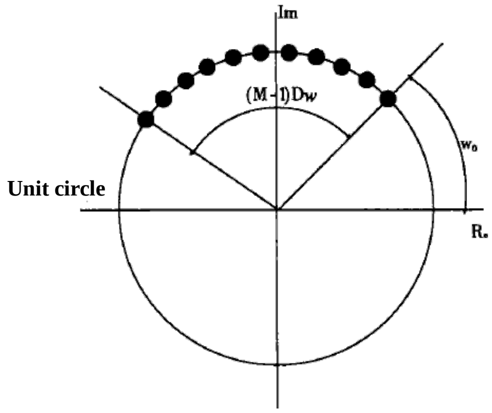

Fig.1 Linear Frequency Modulation Transform frequency samples

$$\omega_k = \omega_0 + k\Delta\omega\ ,\quad k = 0,1,\dots,M-1 \tag{1}$$

The start frequency increment W can be arbitrarily selected. For the special case of DFT, where$\omega_0 = 0$, $M=N$ and$\Delta\omega = 2\pi/N$. The Fourier transform corresponding to this normal set of frequency samples is:

$$X(e^{iw_k}) = \sum_{k=0}^{n-1} x[n] e^{-iw_k n}\quad k = 0,1,\dots,M-1 \tag{2}$$

The calculation process of CZT is generally transformed into linear convolution and then calculated by FFT. Because of the fast computation, CZT has not only been widely used in spectrum refinement, but also used in other algorithms that require fractional Fourier transform. The M-point Chirp Z transform is widely used to refine the spectrum. From the formula point of view, it can obtain a sufficiently high accuracy. But in fact, its resolution is also limited, which usually related to the length of the data, the interception of the signal, etc.

2.2 Definition of CZT.

Let *X(n)* be a known time signal whose *Z* transform is $X(z) = \sum_{k=0}^{+\infty} x[n] Z^{-n}$, where $Z = e^{STs} = e^{(\sigma + j\Omega)\mathrm{Ts}} = Ae^{j\omega}$, with $A=e^{\sigma Ts}$ and $\omega = \Omega\mathrm{Ts}$.

Let $Z_t = \mathrm{A}\omega^{-\mathrm{r}}$, $A = A_0 e^{j\theta 0}, \omega = \omega_0 e^{-j\varphi 0}$ then, $Z_t = A_0 e^{j\theta 0} \omega_0{}^{-r} e^{j\varphi 0 r}$, with A_0 , ω_0 are arbitrary positive real number. When *r=0, 1, 2, ⋯⋯, ∞,* we can get the points $Z_0, Z_1, Z_2, \dots\dots, Z_\infty$ on the *Z* plane, and take the *Z* transform at these points. Then

$$X(Zr) = CZT[x(n)] = \sum_{n=0}^{+\infty} x[n] Zr^{-n} = \sum_{n=0}^{+\infty} x(n) A^{-n} \omega^{nr} \tag{3}$$

When *r=0,* $Z_0 = A_0 e^{j\theta 0}$, this point is in the *Z* plane, and the argument angle is $\theta 0$, which is the starting point of CZT. When it changes with *r*, points $Z_0, Z_1, Z_2, \dots\dots, Z_\infty$ constitute the path of CZT transformation.

3. Countermeasures for the Combination with Production Practice

3.1 Production practice.

The detection of piano pitch requires much higher pitch detection than normal Mandarin tone, as mentioned above, where one note of the first string is 0.02 Hz. So the first requirement is accuracy. The second requirement is anti-interference. The harsh production environment in the factory is far from what most people can imagine. Because it is an assembly line operation, in the front line is the process of planing, milling, drilling, and the sound is tuning at the back of the assembly line. The working environment of the tuning workers is a small shed made of thin boards. The ability to remove noise disturbances has become a key to the success of the project. And the third requirement is real-time. Because of the large output in the factory, the efficiency requirements are extremely high.

It was stipulated that a piano must be adjusted within 15 minutes and the time spent on each string was only 10 seconds. The fourth requirement is stability. The program must be able to measure the pitch after the worker has struck the string, but it cannot be measured sometimes, therefore it must be ensured that workers can see the results every time they look up.

3.2 Countermeasures.

The note score differs between the first string and the last string, and the difference is very large. Setting the accuracy value to be the same is not effective in practice. Therefore, the precision value can be set as dynamic, and one note value of the current string is used as the accuracy value requirement. For the requirements of anti-interference, the screen points can be obtained by using the data values obtained after CZT transformation, as shown in Figure 2.

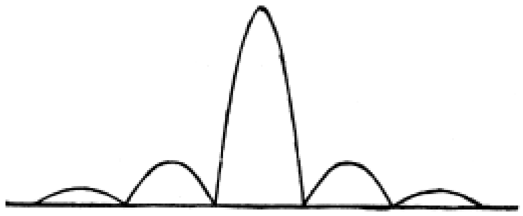

Fig.2 Theoretical transformation graph

t is easy to confirm that the highest point is the corresponding point of the pitch frequency, but in practice it is a continuous frequency measurement. When the noise of the tone is relatively small, the calculated value is not shown as shown in Figure 2, but as shown in Figure 3.

Fig.3 Actual transformation graph

For this case, it cannot simply be assumed that the highest point is the corresponding point of the pitch frequency. It is not feasible to simply set a threshold to separate the above two types of graphs. Even setting a threshold based on each string is not feasible. Because various factors such as the selection of the sampling time period, the strength of the sound, etc. will result in a large range of variation between the calculated maximum values. To solve this problem, onel solution is to analyze the above two pictures, it can be found that when the audio signal is detected, the spectrum has a very high peak relative to the other wave peak, while the frequency spectrum without the audio signal also has the peak, but the wave peak is not so high as the other wave peaks. From this point of view, we can find a way to distinguish these two kinds of graphs. And based on the actual test situation, it can be found the matching accuracy is high enough, and will not interfered by noise level.

The solution to real-time problems must be done with the assurance of accuracy over time. Obviously, it cannot be measured only once after the string is struck, because the frequency will gradually change with the vibration of the string, so it must be detected in real time, and the sampling time must be as short as possible. Firstly, according to the accuracy requirements, the length of sampling time should be calculated theoretically, and at the same time, woker's factor should be also considered: If the sampling time is too long, workers will feel inefficiency when they look up at the results, and some research results in this area show that the waiting time cannot exceed 0.2s. Secondly, Combining repeated tests with the actual conditions in the factory, to determine the reasonable sampling period, so that the calculated frequency value can reflect changes in the pitch caused by the vibration of the string in real time, and at the same time meet the worker's requirements for efficiency.

4. The Principle and Process of the New Pitch Detection Device

The pitch detection device is very simple: only a normal sound card and a microphone are needed. The main work is done by software, and the detection process is shown in Figure 4.

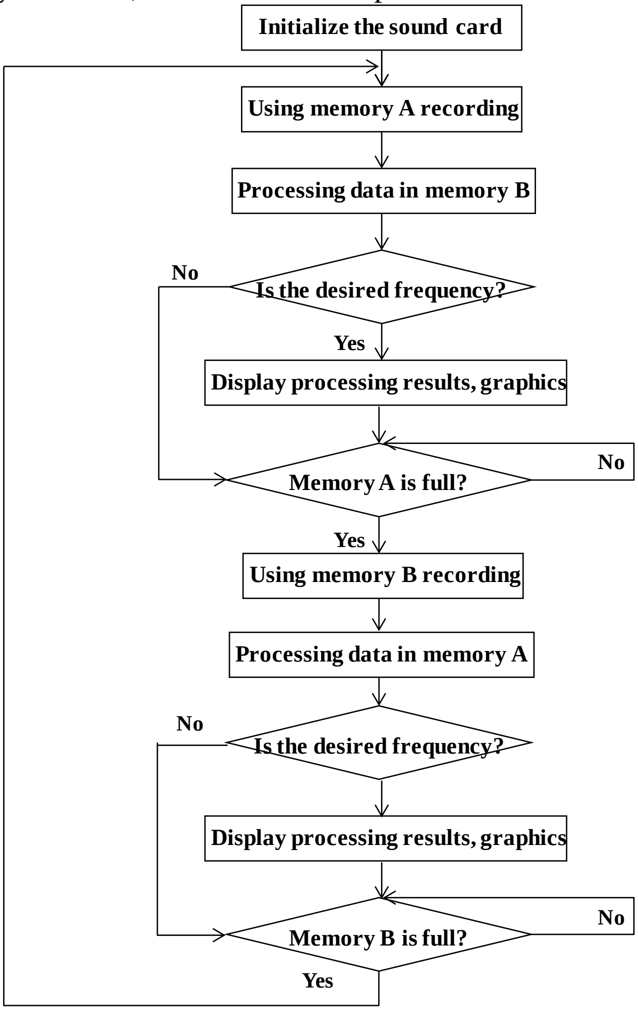

Fig.4 Detection flow chart

This program is written using VC6.0. The judgments in the above flowchart are all implemented by callback functions. With multi-threading, it is also possible to handle multiple tasks at the same time, such as the worker controlling the detection process with a foot switch.

5. Summary

CZT is widely used for spectrum refinement work. Obviously, CZT can be used to make a relatively accurate frequency estimation of a single frequency, and the simulation also proves this point. This resolution can be very high. When using a general Fourier transform, spectral lines at two adjacent discrete points give a ambiguous meaning or have two adjacent frequency components, or have a large frequency component between them. At this time, using CZT will get good results and give a clear explanation. The completion of the project solved a specific and complex practical problem for the musical instrument factory, which greatly increased the production efficiency of the factory and achieved remarkable economic benefits. It should also be noted that if the database in the software is changed, the software can also be used for pitch detection of guitars, zithers, and other musical instruments.

References

[1] Proak is J D, M anolak is D G. Digital Signal Processing [M]. Beijing: Publishing House of

Electronics Industry, 2007: 402-407.

[2] Qian Ke mao, Li Chuanqi, New spectrum correction method based on chirp Z transform [J]. Journal of Vibration Engineering, 2010, 13 (4): 628-631.

[3] Hu Guangshu. Digital Signal Processing -Theory, Algorithms, and Implementation. Beijing: Tsinghua University Press, 2014 (08): 153-154.

[4] Oppenheim A V, Schafer RW, Buck JR. Discrete-time signal processing [M]. Xian: Xian Jiaotong University Press, 2010: 528-532.